Mo et al. BioMedical Engineering OnLine (2025) 24:39
https://doi.org/10.1186/s12938-025-01360-1

BioMedical Engineering OnLine

RESEARCH Open Access

Forecasting motion trajectories of elbow and knee joints during infant crawling based on long–short-term memory (LSTM) networks

Jieyi Mo[1], Qiliang Xiong[1*], Ying Chen[1], Yuan Liu[3], Xiaoying Wu[2], Nong Xiao[3] and Wensheng Hou[2]

*Correspondence: 70898@nchu.edu.cn

[1] Department of Biomedical Engineering, Nanchang Hangkong University, Jiangxi, China
[2] School of Bioengineering, Chongqing University, Chongqing, China
[3] Department of Rehabilitation, Children's Hospital of Chongqing Medical University, Chongqing, China

Abstract

Background: Hands-and-knees crawling is a promising rehabilitation intervention for infants with motor impairments, while research on assistive crawling devices for rehabilitation training was still in its early stages. In particular, precisely generating motion trajectories is a prerequisite to controlling exoskeleton assistive devices, and deep learning-based prediction algorithms, such as Long–Short-Term Memory (LSTM) networks, have proven effective in forecasting joint trajectories of gait. Despite this, no previous studies have focused on forecasting the more variable and complex trajectories of infant crawling. Therefore, this paper aims to explore the feasibility of using LSTM networks to predict crawling trajectories, thereby advancing our understanding of how to actively control crawling rehabilitation training robots.

Methods: We collected joint trajectory data from 20 healthy infants (11 males and 9 females, aged 8–15 months) as they crawled on hands and knees. This study implemented LSTM networks to forecast bilateral elbow and knee trajectories based on corresponding joint angles. The data set comprised 58, 782 time steps, each containing 4 joint angles. We partitioned the data set into 70% for training and 30% for testing to evaluate predictive performance. We investigated a total of 24 combinations of input and output time-frames, with window sizes for input vectors ranging from 10, 15, 20, 30, 40, 50, 70, and 100 time steps, and output vectors from 5, 10, and 15 steps. Evaluation metrics included Mean Absolute Error (MAE), Mean Squared Error (MSE), and Correlation Coefficient (CC) to assess prediction accuracy.

Results: The results indicate that across various input–output windows, the MAE for elbow joints ranged from 0.280 to 4.976°, MSE ranged from 0.203° to 59.186°, and CC ranged from 89.977% to 99.959%. For knee joints, MAE ranged from 0.277 to 4.262°, MSE from 0.229 to 53.272°, and CC from 89.454% to 99.944%. Results also show that smaller output window sizes lead to lower prediction errors. As expected, the LSTM predicting 5 output time steps has the lowest average error, while the LSTM predicting 15 time steps has the highest average error. In addition, variations in input window size had a minimal impact on average error when the output window size was fixed. Overall, the optimal performance for both elbow and knee joints was observed with input–output window sizes of 30 and 5 time steps, respectively, yielding an MAE of 0.295°, MSE of 0.260°, and CC of 99.938%.

© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.

Conclusions: This study demonstrates the feasibility of forecasting infant crawling trajectories using LSTM networks, which could potentially integrate with exoskeleton control systems. It experimentally explores how different input and output time-frames affect prediction accuracy and sets the stage for future research focused on optimizing models and developing effective control strategies to improve assistive crawling devices.

Keywords: Trajectory prediction, Infant crawling, LSTM networks

Introduction

Before infants achieve independent walking, developmental milestones include rolling over, sitting, and crawling on hands and knees. Among these, crawling represents the first gross motor behavior involving the coordination of elbow and knee joints [1]. Specifically, research shows that increased frequency and duration of crawling help develop motor skills and support the acquisition of walking abilities [2, 3]. Conversely, insufficient crawling experience can lead to abnormal gait patterns [4]. It has been suggested that crawling training has been found to benefit motor function rehabilitation and cognitive development, particularly in children with cerebral palsy and balance impairments related to stroke [5–7]. For example, crawling supports motor function development in infants with delays and enhances cerebellar motor stability [8]. Extended crawling practice also stimulates the neuromuscular system, aiding in the recovery and rebuilding of neuromuscular functions and improving overall rehabilitation outcomes [9]. Given its benefits, hands-and-knees crawling is gaining attention as a promising rehabilitation approach for infants with motor impairments, leading to increased interest in developing assistive devices for crawling training.

Several passive-guided exoskeleton devices have emerged to assist patients in crawling training. For instance, FITCRAWL in Australia has developed a crawling robot designed for physical exercise in healthy adults [10]. Ghazi et al. developed an assistive crawling device for children with cerebral palsy, using EEG-based neuroimaging and a custom wearable motion capture system to monitor development [11]. In addition, Jiang et al. focused on coordinating hand and knee movements in typical infant crawling to design a new rehabilitation aid for cerebral palsy, incorporating an assisted crawling training apparatus [12]. However, such methods that control based on predefined movement patterns overlook the initiative and proactivity of infant crawling movement. They may lead to issues such as motion coordination problems and dragging of the wearer, hindering the recovery of the patient's motor functions. Accurately predicting the future trajectory of infant crawling could improve the performance of rehabilitation devices by adding a feedforward control component. This would allow the device to better adapt to changes in crawling patterns, synchronize more smoothly with the user's movements, and reduce disruptions when the user alters their motion. Therefore, forecasting crawling trajectories is essential for developing effective motion planners and high-level controllers for exoskeleton crawling devices.

Despite the potential benefits, research on predicting infant crawling trajectories is still limited. This study makes a significant contribution by being the first to apply deep learning techniques to predict crawling trajectories in infants. Specifically, long–short-term memory (LSTM) networks [13], a recent advancement in time series prediction, are well-suited for this task. Since trajectory data exhibit temporal correlations, LSTM networks are ideal for modeling the non-linear, dynamic behavior of movement patterns, enabling accurate predictions of future positions based on past sequences [14–16]. Accordingly, this paper aims to evaluate the feasibility of using LSTM networks to predict infant crawling trajectories with high accuracy.

Given that the elbow and knee dominate the rhythmical flexion and extension of limbs during crawling on hands and knees, three-dimensional trajectory data of these two joints and the corresponding joint angle were calculated when infants were crawling at their self-selected velocity. Then, an LSTM autoencoder model was used to predict the angle of elbow and knee motion variables, exploring the feasibility of accurately predicting infant crawling trajectories. In addition, we explored the influence of input and output window lengths on prediction accuracy and put forward technical recommendations. The remainder of the paper is structured as follows: "Related works" section reviews related work on trajectory prediction using deep learning methods, with a focus on LSTM networks. "Methods" section outlines the data collection protocol, data preprocessing, and implementation details of the deep learning model. "Results" section presents the results, while "Discussion" section discusses the implications, limitations, and future directions. Finally, "Conclusions" section concludes the paper.

Related works

Recent advances in time series prediction have highlighted the effectiveness of deep learning methods for forecasting movement trajectories. LSTMs are particularly advantageous due to their ability to learn from sequential data and maintain long-term dependencies, allowing them to use past motion patterns to make accurate predictions about future movements [17]. Several studies have applied LSTMs to predict gait trajectories (an overview is provided in Table 1). For example, Liu et al. developed a deep spatiotemporal model consisting of LSTM units to forecast the next two time steps, smoothing predictions by averaging them [18]. Zaroug et al. implemented an autoencoder LSTM to predict trajectories of linear acceleration and angular velocity [19]. They experimented with input time steps ranging from 5 to 40 time steps to predict future trajectories over 5 or 10 time steps (equivalent to 30 ms or 60 ms). Su et al. proposed an LSTM with a weighted discount loss function to predict angular velocities of the thigh, calf, and foot segments [20]. They used 10 or 30 time steps as input to predict future trajectories over 5 or 10 steps, corresponding to 100 ms and 200 ms, respectively. Hernandez et al. utilized a hybrid convolutional neural network (CNN) and LSTM neural network, DeepConvLSTM, to predict motion trajectories with an average Mean Absolute Error (MAE) of 3.6° [21]. Jia et al. employed LSTM units combined with a feature fusion layer that integrates kinematic (joint angles) and physiological (electromyography) data for trajectory prediction [22]. Zarough et al. also compared vanilla LSTM, stacked LSTM, bidirectional LSTM, and autoencoder LSTM [23], while Zhu et al. used attention-based CNN–LSTM to forecast trajectories over the next 60 ms [24]. Other notable studies include Challa et al., who proposed an LSTM-based human gait trajectory generator using data collected from Microsoft Kinect V2 [25], and Semwal et al., who introduced an LSTM–CNN sequential model capable of generating stable gait trajectories within a speed range of 0.49–1.76 m/s, achieving a high correlation of 0.98

Table 1 Overview of recently related work about trajectory prediction using LSTM

Tasks	Reference	Methodology	Results	Publication date
Walking	Zaroug et al. [19]	Encoder–decoder LSTM	Correlation in the order of 0.98 between predicted and actual trajectory	2020
Walking	Su et al. [20]	LSTM	A correlation of 0.98 in the predicted trajectory and 95% accuracy in phase prediction	2020
Walking & running	Hernandez et al. [21]	DeepConvLSTM	MAE in range 2.2(0.9)–5.1(2.7) degrees	2021
Walking	Jia et al. [22]	LSTM	RMSE in rage 0.348–0.713 degrees, correlation in the order of 0.99	2021
Walking	Zarough et al. [23]	LSTM	NRMSE in range 2.82–5.31%	2021
Walking	Zhu et al. [24]	Attention-based CNN–LSTM	Within a predicted horizon of 60 ms, the prediction RMSE is as low as 0.317 degrees	2021
Walking	Challa et al. [25]	LSTM	The gait trajectories obtained through the proposed model are also validated on the HOAP-2 robot simulator	2022
Walking	Semwal et al. [26]	LSTM–CNN	A high correlation of 0.98 between the actual and the predicted trajectories, and an R-2 Score of 0.94 is obtained	2023
Walking	Romero-Sorozábal, et al. [27]	LSTM & Regression	RMSE of 13.40 mm and a correlation coefficient of 0.92 for the regression model, and RMSE of 12.57 mm and a correlation of 0.99 for the LSTM	2024

between actual and predicted trajectories, along with an R-squared score of 0.94 [26]. In addition, Romero-Sorozábal et al. presented regression and LSTM models for predicting three-dimensional trajectories [27].

It is important to note that previous studies have focused on predicting limb movement trajectories during human walking, achieving promising results using LSTM models. However, unlike walking, the limb trajectories during infant crawling exhibit greater variability, which complicates prediction and raises concerns about the feasibility of making accurate forecasts. Therefore, the main contributions of this paper are three-folds. First, we provide a comprehensive theoretical overview of the significance of infant crawling, the rehabilitative benefits of crawling training devices, and the approach to actively controlling these devices using deep learning algorithms. Second, we assess the performance of the LSTM network in forecasting infant crawling trajectories, presenting detailed prediction results for the first time. Finally, we examine how the length of input and output windows impacts prediction accuracy and offer technical recommendations.

Results

LSTM network performance for varying input and output window sizes

The LSTM model was trained using 24 combinations of input and output window sizes. Input window sizes ranged from 10, 15, 20, 30, 40, 50, 70, and 100 time steps, and output window sizes were 5, 10, and 15 time steps. The following results show the model's performance in terms of mean absolute error (MAE), mean square error (MSE), and correlation coefficient (CC).

As shown in Fig. 1, when the output window was fixed at five time steps, the MAE for all four joints ranged from 0.295 to 0.382°, the MSE from 0.260 to 0.430°, and the CC from 99.915% to 99.941%. Taken together, the overall optimal performance was achieved with an input window size of 30 time steps (MAE = 0.295°, MSE = 0.260°, CC = 99.938%) when the output window was fixed at five time steps. This trend was also observed when the output window sizes were 10 or 15 time steps.

Figures 2, 3, 4 further show that the output window size has a notable impact on prediction accuracy, with smaller windows generally resulting in lower errors. Specifically, the five-time-step output window produced the lowest average error, whereas the 15-time-step window resulted in the highest error. Accordingly, in the following section, we will analyze the performance of models across specific joints, examining varying input window sizes using a fixed output window of five time steps, as well

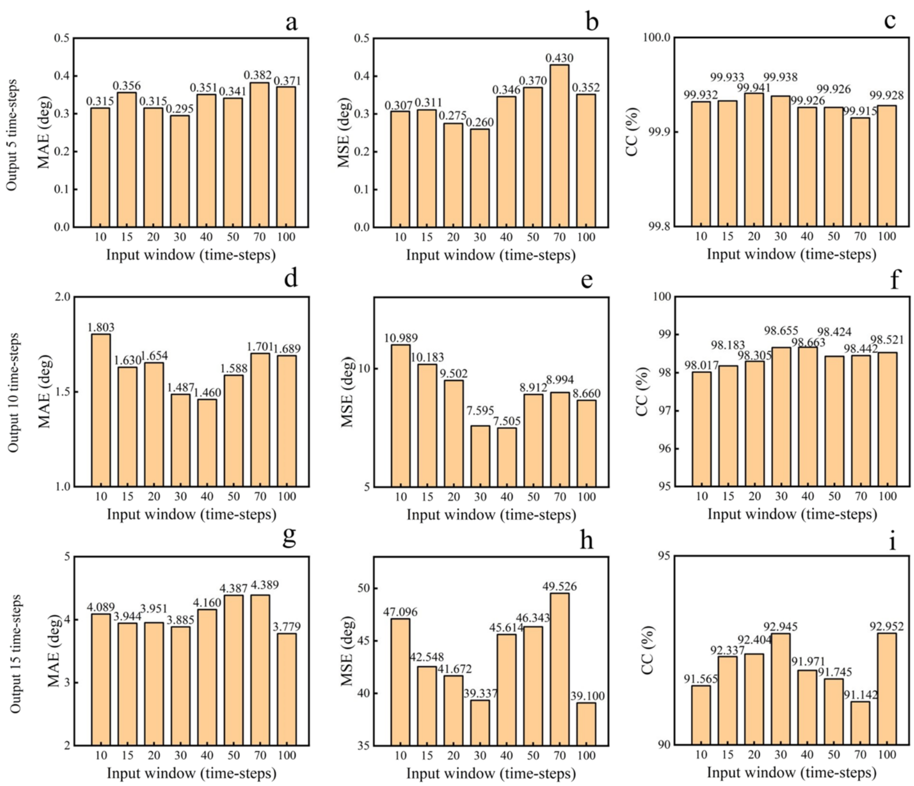

Fig. 1 LSTM model's performance was assessed using MAE, MSE, and CC across different input window sizes, ranging from 10 to 100 time steps. The output window sizes were set to 5 (**a**–**c**), 10 (**d**–**f**), and 15 time steps (**g**–**j**). The bar chart presents the average performance metrics for the bilateral elbow and knee joints, illustrating the effects of different input window sizes

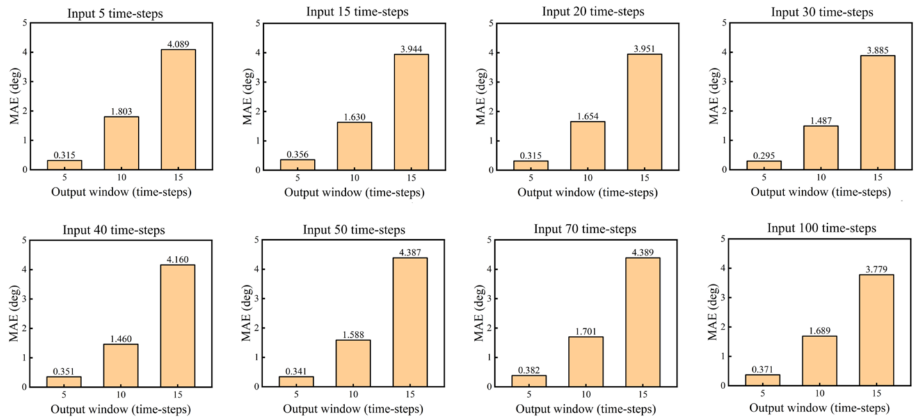

Fig. 2 LSTM model's performance was evaluated using MAE across various output window sizes of 5, 10, and 15 time steps. Input window sizes ranged from 10 to 100 time steps, including 10, 15, 20, 30, 40, 50, 70, and 100. The bar chart displays the average performance metrics for the bilateral elbow and knee joints, highlighting the impact of different output window sizes

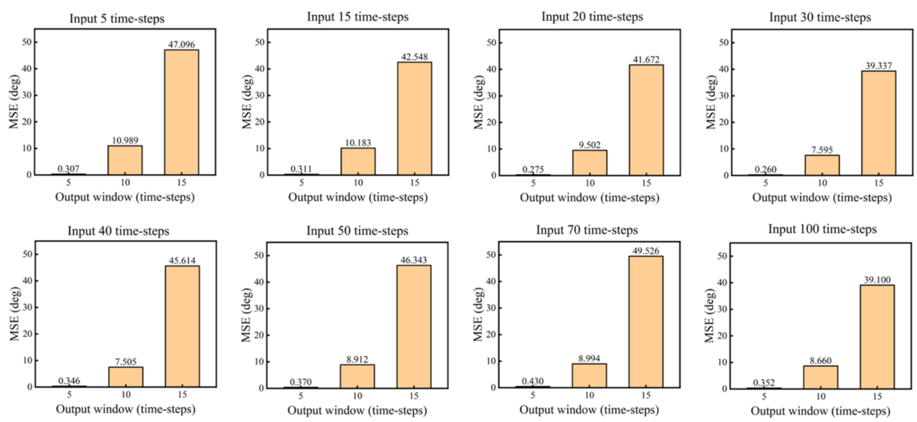

Fig. 3 LSTM model's performance was evaluated using MSE across various output window sizes of 5, 10, and 15 time steps. Input window sizes ranged from 10 to 100 time steps, including 10, 15, 20, 30, 40, 50, 70, and 100. The bar chart displays the average performance metrics for the bilateral elbow and knee joints, highlighting the impact of different output window sizes

as models with different output window sizes using a fixed input window of 30 time steps.

The performance of models with a fixed output window of five time steps

We assessed the impact of eight different input window sizes—10, 15, 20, 30, 40, 50, 70, and 100 time steps—on the model's prediction accuracy, with the output window fixed at five time steps. Figure 5 illustrates how different input window sizes influence the model's performance across specific joints, with smaller errors indicating higher accuracy. The performance metrics for the left elbow (LElbow), right elbow (RElbow), left knee (LKnee), and right knee (RKnee) are detailed in Fig. 11a–d. Our analysis indicates that the input window size of 30 time steps produced the most accurate predictions overall.

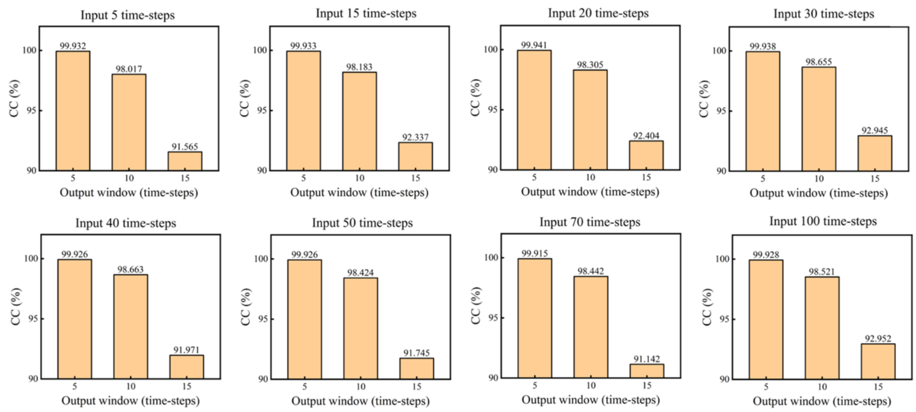

Fig. 4 LSTM model's performance was evaluated using CC across various output window sizes of 5, 10, and 15 time steps. Input window sizes ranged from 10 to 100 time steps, including 10, 15, 20, 30, 40, 50, 70, and 100. The bar chart displays the average performance metrics for the bilateral elbow and knee joints, highlighting the impact of different output window sizes

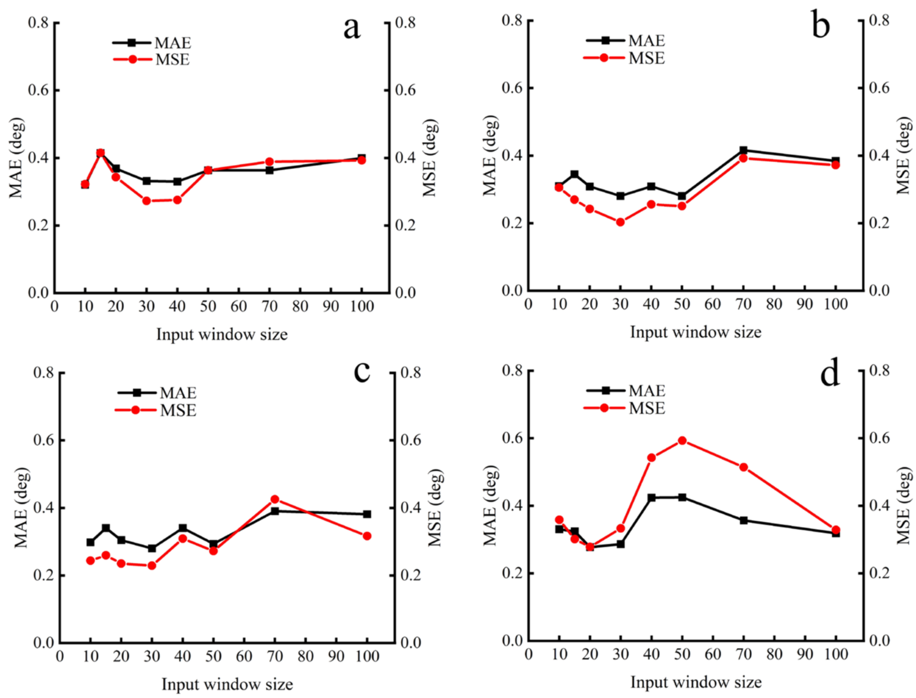

Fig. 5 With an output window fixed at five time steps, the model's performance varies across different input window sizes. MAE is represented in black, and MSE in red. **a** Left elbow joint. **b** Right elbow joint. **c** Left knee joint. **d** Right knee joint

The performance of models with a fixed input window of 30 time steps

We examined the impact of output window sizes set to 5, 10, and 15 time steps on the model's performance, with the input window fixed at 30 time steps. Smaller prediction errors indicate better accuracy. The performance metrics for the left elbow (LElbow), right elbow (RElbow), left knee (LKnee), and right knee (RKnee) are detailed in Fig. 6a–d, the results reveal that the model's performance varied significantly with different

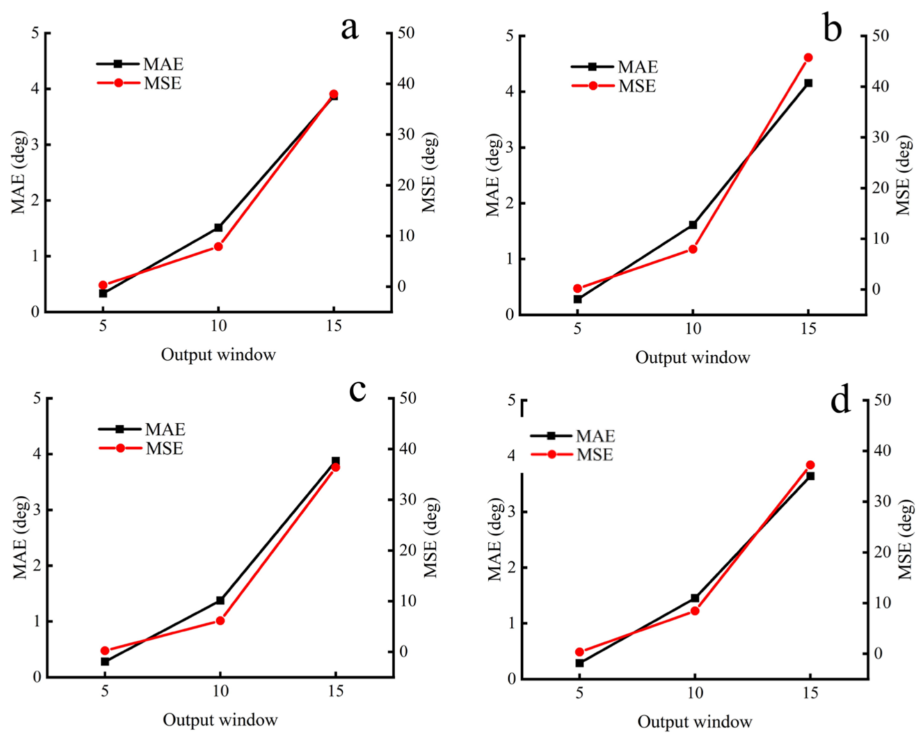

Fig. 6 With an input window fixed at 30 time steps, the model's performance varies across different output window sizes. MAE is represented in black, and MSE in red. **a** Left elbow joint. **b** Right elbow joint. **c** Left knee joint. **d** Right knee joint

output window sizes. Larger output windows were associated with higher prediction errors, which aligns with our previous findings.

The joint trajectories predicted with an input window of 30 time steps and an output window of 5 time steps

Figure 7 illustrates the optimal model's predictions for four joints throughout a complete cycle. The optimal model employs a sliding window with an input size of 30 time steps and an output size of 5 time steps. Performance metrics for the left elbow (LElbow), right elbow (RElbow), left knee (LKnee), and right knee (RKnee) are summarized in Table 2. These results indicate that the LSTM model effectively forecasts joint trajectories, achieving an average MAE of 0.295°, an average MSE of 0.260°, and an average CC of 99.938%. The best performance was observed with input and output window sizes of 30 and 5 time steps, respectively.

Discussion

In this study, our objective was to develop and evaluate an LSTM autoencoder model to predict trajectories of 4 motion variables (Y_1,Y_2,Y_3,Y_4), exploring the feasibility of accurately predicting infant crawling trajectories. To our knowledge, this research represents the first application of deep learning models to predict crawling trajectories in infants. LSTM, a type of gated recurrent network, was chosen for this task due to its proven success in handling sequential data [23]. The key advantage of LSTM is

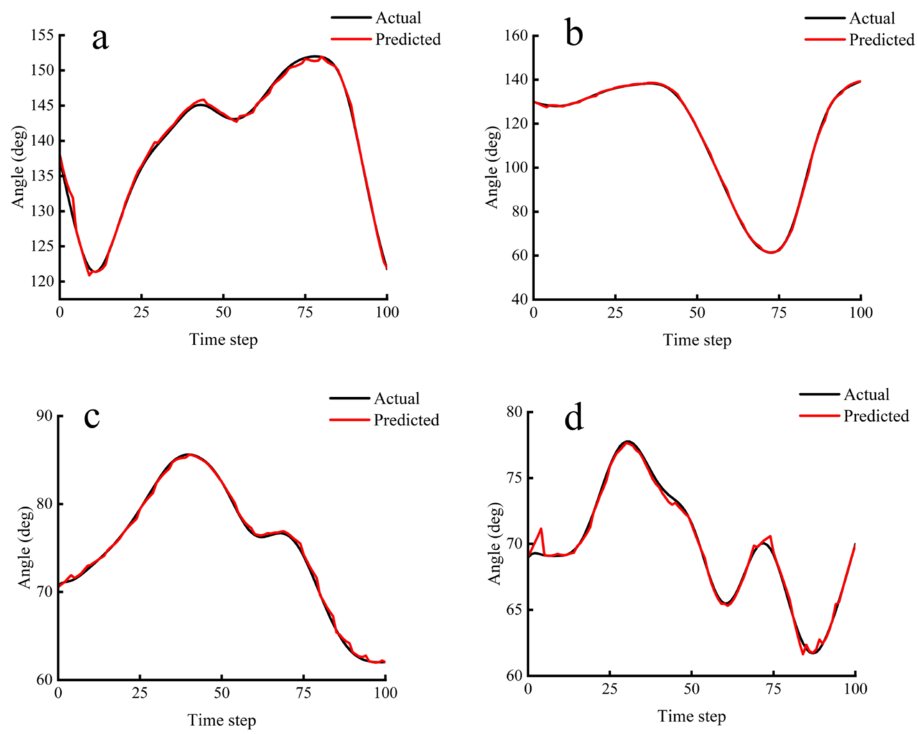

Fig. 7 Joint trajectories were predicted with an input window of 30 time steps and an output window of 5 time steps. The figure displays predicted trajectories (in red) and actual trajectories (in black) for (**a**) left elbow joint angle, (**b**) right elbow joint angle, (**c**) left knee joint angle, and (**d**) right knee joint angle

Table 2 Model's performance metrics across specific joints using an input window of 30 time steps and an output window of 5 time steps

Joints	MAE (deg)	MSE(deg)	CC (%)
LElbow	0.332	0.273	99.935
RElbow	0.280	0.203	99.959
LKnee	0.280	0.229	99.935
RKnee	0.287	0.333	99.923
Average	0.295	0.260	99.938

its ability to account for the order of values in input sequences, enabling it to learn long-term dependencies [23]. Our results demonstrate that LSTM models can effectively predict changes in elbow and knee joint angles (e.g., Fig. 7). The optimal performance for both joints was achieved with input–output window sizes of 30 and 5 time steps, respectively, resulting in an MAE of 0.295°, an MSE of 0.260°, and a correlation coefficient (CC) of 99.938%. These findings suggest that incorporating LSTM-based predictions into assistive device controllers could improve their functionality by adding a feedforward component, thus reducing dependence on feedback mechanisms [28]. This integration would allow assistive devices to better adapt to changes in crawling patterns, enhancing alignment with the user's intent and minimizing interruptions during movement transitions [29–32]. Furthermore, predicting future trajectories could help monitor the risk of imbalance and falls, facilitating early intervention through remote alerts [33–37].

In assistive device control systems, it is essential to strike a balance between prediction accuracy and processing speed. The input window should be large enough to ensure reliable predictions but not so large that it slows down the system. While previous research by Banos et al. recommended a 1–2 s window for human activity recognition [38], no specific guidelines exist for predicting infant motion trajectories. To fill this gap, we tested various input window sizes to determine the optimal predictive model. Initially, we varied the input window between 10 and 100 time steps, keeping the output window fixed at 5 time steps. As shown in Fig. 5, the MAE for elbow joints ranged from 0.295° (with a 30-time-step input window) to 0.382° (with a 70-time-step input window). Similarly, MSE ranged from 0.260° (30-time-step input window) to 0.430° (70-time-step input window), and the correlation coefficient (CC) ranged from 99.915% (70-time-step input window) to 99.941% (20-time-step input window). These results show that the input window size had minimal impact on LSTM model accuracy, with an optimal input window of around 30 time steps and a poorer performance observed with a 70-time-step input window. This aligns with existing literature indicating that prediction errors increase when the input window exceeds 30 time steps [23]. Subsequently, as shown in Fig. 1, we further tested different output window sizes by varying the input window between 10 and 100 time steps while fixing the output window at 10 and 15 time steps. The results confirmed that the optimal input window remained 30 time steps, but a 5-time-step output window provided better performance, as evidenced by lower MAE (0.295°), MSE (0.260°), and higher CC (99.941%) compared to the 10-time-step and 15-time-step output windows. This contrasts with findings by Kolaghassi et al., which suggested that longer input windows reduce prediction errors when the output window exceeds 12 time steps [39].

In addition, we assessed the impact of output window size on model accuracy. With a fixed input window of 30 time steps, varying the output window size from 5 to 15 time steps revealed a significant decline in performance as the output window increased, evidenced by a rise in MAE and MSE for both the elbow and knee joints (Fig. 6). We observed that model performance declined significantly as the output window increased, as evidenced by a marked rise in MAE and MSE for the bilateral elbow and knee joints. These results support our previous findings and suggest that output windows larger than five time steps may not be reliable for predicting crawling trajectories. To address this challenge, alternative deep learning models, such as bidirectional LSTM [40], or hybrid approaches could offer potential solutions for improving predictive accuracy over longer time frames.

Several limitations are acknowledged in this study. First, the differences between actual and predicted trajectories (particularly those shown in Fig. 7d) are more significant than the mean absolute error suggests. It is difficult to determine whether this discrepancy is due to the model's inability to generalize certain crawling patterns or potential issues with the sample data, such as sensor inaccuracies or labeling errors. Data collection in infants is inherently challenging, leading to a limited data set. Although the current study includes 20 healthy infants, this small sample size may affect the generalizability of the results. To improve the model's reliability and applicability, we recommend expanding the data set to include a more diverse population, especially infants with motor impairments. A larger and more varied data set would help ensure that the LSTM model generalizes well to a broader range of crawling behaviors, including both typically developing infants and those undergoing rehabilitation. Second, crawling patterns vary significantly among infants, particularly in those with motor impairments. This study primarily focuses on healthy infants, whose crawling trajectories may not fully reflect the complexity and variability seen in infants with developmental delays or physical disabilities. Future research should include a broader spectrum of crawling behaviors, particularly those affected by conditions, such as cerebral palsy or other motor impairments. Expanding participant diversity would improve the model's ability to predict trajectories beyond five time steps, supported by a more comprehensive validation set. This approach would also help optimize training epochs and reduce the risk of overfitting [41]. Finally, while the use of LSTM networks for trajectory prediction in controlled settings shows promise, several challenges remain when applying them to real-world scenarios, particularly in controlling assistive crawling devices. Real-world environments involve dynamic factors, such as uneven surfaces, obstacles, or external disturbances, which can significantly alter crawling patterns. In addition, changes in an infant's posture, fatigue, or motivation during rehabilitation may further complicate prediction. To address these challenges, future research should focus on integrating real-time sensor data and adaptive algorithms to ensure that the system remains robust and responsive in real-world settings. Moreover, designing an assistive device that can adjust to fluctuations in crawling behavior is essential for effective rehabilitation.

Conclusions

In summary, this study designed a framework for predicting infant crawling motion trajectories using an LSTM network, confirming the feasibility of predicting joint motion trajectories during infant crawling. In addition, we explored various input and output window sizes to quantify how performance is influenced by input data volume and future horizon length. The experimental results show that the LSTM model can accurately predict the elbow and knee trajectory with an average mean square error (MAE $=$ 0.295°, MSE $=$ 0.260°, CC $=$ 99.938%), while the optimal performance was observed with input–output window sizes of 30 and 5 time steps, respectively. A potential application of our method is in the control of crawling rehabilitation devices, where predicted model trajectories can serve as proxies for user intent. These intents can be integrated into the control hierarchy of exoskeletons, particularly in high-level control, which detects user intentions and passes them to lower levels to generate appropriate motion commands, potentially enhancing clinical rehabilitation outcomes for infants with conditions like cerebral palsy.

Methods

Participants

Twenty healthy infants (11 males and 9 females, aged 8–15 months) were recruited from local child health clinics. All participants were full-term births and had no reported neurological impairments during the neonatal period, as confirmed by their parents [42]. Kinematic data were captured using a motion capture system (Raptor-E, Motion Analysis

Corporation, USA) with six high-speed digital cameras operating at 100 frames per second. During recordings, infants wore only diapers, and reflective markers were placed on specific anatomical landmarks: shoulders (lateral to the acromion), elbows (lateral epicondyle), wrists (ulnar styloid process), hips (posterior superior iliac spine), knees (lateral joint line), ankles (lateral malleolus), and trunk (shoulder blade).

Before data collection, infants had a warm-up period on a crawling mat measuring 360 cm × 120 cm. They were encouraged to crawl toward toys or in response to their mother's calls (as shown in Fig. 8). A valid trial was defined as a continuous sequence of at least three complete and consecutive strides. Only straight crawling sequences without interruptions or deviations were included in the analysis. The initial and final steps of each sequence were excluded. Crawling cycles were defined based on the landing time of the right wrist joint, resulting in 582 valid cycles. Each cycle was uniformly resampled to cover 0–100% of the crawling cycle, yielding a total of 58, 782 time steps for analysis.

The experiments were conducted at the Department of Rehabilitation Center, Children's Hospital of Chongqing Medical University. The study was approved by the hospital's Ethics Committee (Approval number: 065/2011), and informed written consent was obtained from the parents or legal guardians of all participating infants.

Data processing

Given that the elbow and knee dominate the rhythmical flexion and extension of limbs during crawling on hands and knees, In the current study, joint angles of elbow and knee were calculated primarily using three-dimensional coordinate data of adjacent joints in space (displacement in x, y, and z directions). For instance, the elbow joint angle is the angle formed by the lines connecting the wrist, elbow, and shoulder joints. Similarly, the knee joint angle is determined by the lines connecting the hip, knee, and ankle joints.

As depicted in Fig. 9, we constructed spatial vectors to determine these angles. For the elbow joint, we used the coordinates of the shoulder joint (S_x,S_y,S_z), elbow joint (E_x,E_y,E_z), and wrist joint (W_x,W_y,W_z) to form vectors in the elbow–wrist direction ($\overline{\mathrm{ES}}$) and elbow–shoulder direction ($\overline{\mathrm{EW}}$):

$$\overline{\mathrm{ES}} = (S_x - E_x, S_y - E_y, S_z - E_z) \tag{1}$$

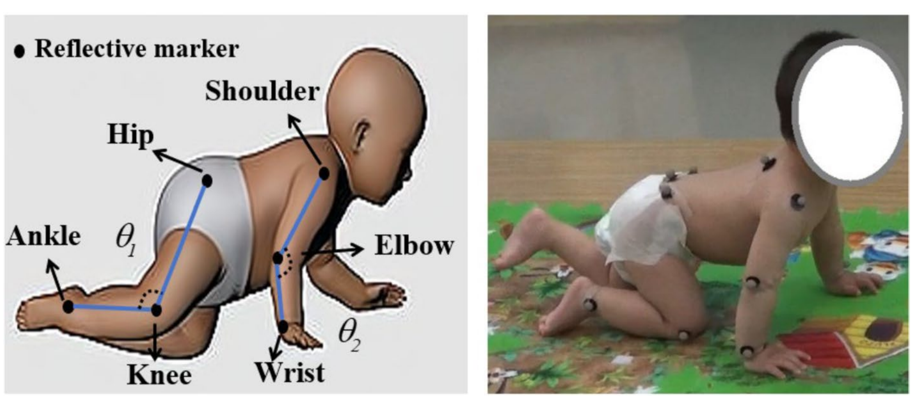

Fig. 8 Placement of the reflective markers and snapshot of data collection

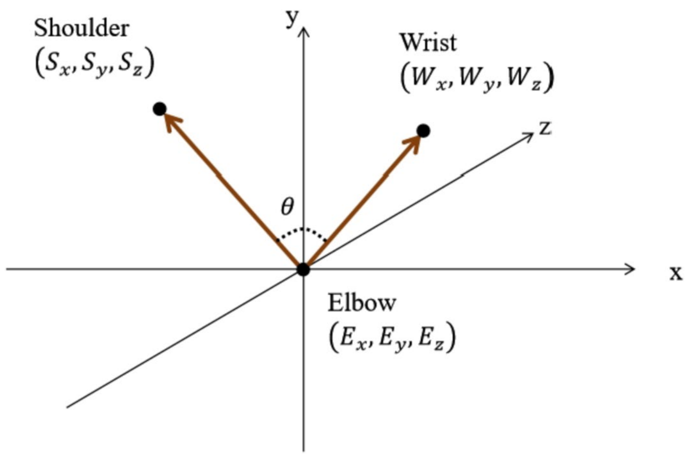

Fig. 9 Illustration of the calculation of elbow joint angles during infant crawling

$$\overline{\mathrm{EW}} = (W_x - E_x, W_y - E_y, W_z - E_z) \tag{2}$$

$$cos\theta = \frac{\overline{\mathrm{ES}} \cdot \overline{\mathrm{EW}}}{|\ \overline{\mathrm{ES}}\ ||\ \overline{\mathrm{EW}}\ |} \tag{3}$$

Accordingly, the calculation of the elbow joint angle can be directly determined by the angle between the spatial vectors $\overline{ES}$ and $\overline{EW}$ as follows:

$$\theta = \arccos \frac{(S_x - E_x) \times (W_x - E_x) + (S_y - E_y) \times (W_y - E_y) + (S_z - E_z) \times (W_z - E_z)}{\sqrt{((S_x - E_x))^2 + (S_y - E_y)^2 + (S_z - E_z)^2} \times \sqrt{(W_x - E_x)^2 + (W_y - E_y)^2 + (W_z - E_z)^2}} \times \frac{180}{\pi} \tag{4}$$

Here, the coordinates (S_x,S_y,S_z) and (E_x,E_y,E_z) represent the positions of the shoulder and elbow joints in three-dimensional space, while (W_x,W_y,W_z) represent the wrist joint's position. This approach ensures a precise determination of joint angles based on their spatial arrangement.

Time series transformation to a supervised learning problem

As we mentioned before, the crawling motion cycles were defined by computing the squared time derivative of the positions (squared of velocity) of the wrist [43], resulting in 582 valid crawling cycles. Each cycle was resampled into 0–100% of the crawling cycle, totaling 58, 782 time steps. Each time step included four joint angles, as shown in Fig. 10, leading to a data set with 58, 782 rows and 4 columns (Y_1, Y_2, Y_3, Y_4). We divided the data set into two parts: 70% for training to optimize model parameters and 30% for testing to assess the model's predictive performance.

The LSTM takes as input 4 parallel feature variables (crawling joint angles) and outputs predictions for the subsequent 4 parallel feature variables (crawling joint angles). As shown in Fig. 11, to accommodate the LSTM's requirement for fixed-length sequences, we applied a sliding window approach to generate these sequences. This approach involves creating input and output windows of fixed length, with the input window providing the data for the model and the output window containing the future predictions.

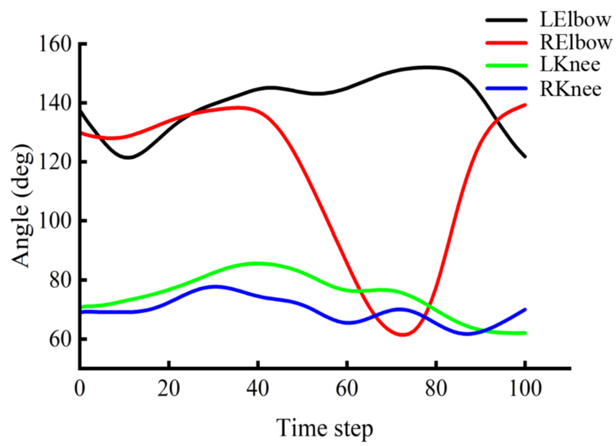

Fig. 10 Typical trajectories of the elbow and knee angles during a single crawling cycle

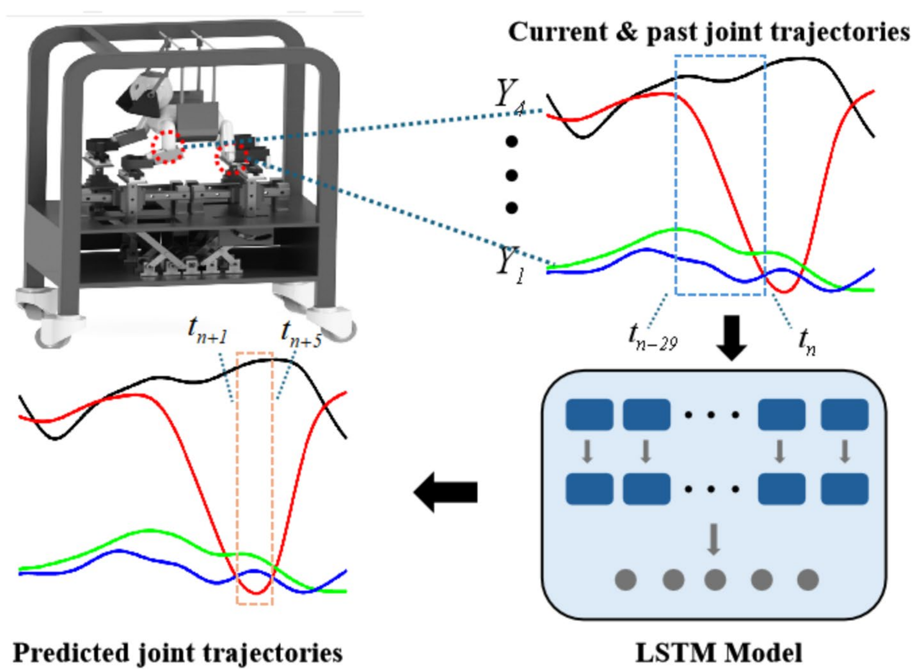

Fig. 11 Schematic diagram of current and past joint angle trajectories to predict future joint angles during infant crawling. The sliding window consists of an input window, an output window, and a sliding step. Both windows comprise a set number of time steps and features

That is, the input window serves as the input data for the LSTM model, while the output window represents the LSTM model's future prediction output. Each input window corresponds to an output window (the target label for training), forming one training sample. The sliding size, which denotes the distance from the start of one sample to the start of the next, always equals the step size of the output window.

In the current study, we used LSTM to forecast elbow and knee trajectories based on varying input and output window sizes. Input window sizes for the LSTM were 10, 15, 20, 30, 40, 50, 70, and 100 time steps (for data captured at a sampling frequency of 100 Hz, these durations correspond to 100, 150, 200, 300, 400, 500, 700, and 1000 ms). The reason for using input window sizes up to 1000 ms is the average length of a crawling cycle for a typically developing infant [42]. This means we trained deep learning models to make predictions based on data from approximately one full crawling cycle, or lower. Output window sizes for the LSTM were 5, 10, and 15 time steps) (corresponding to 50, 100, and 150 ms), allowing us to forecast up to 15% of the crawling cycle.

LSTM neural network

LSTM (long–short-term memory) networks are a specialized type of recurrent neural network designed to address some limitations of traditional models. While conventional recurrent neural networks are effective for processing sequential data, they often struggle with problems, such as gradient vanishing and exploding, which hinder their ability to capture long-term dependencies. LSTM networks enhance traditional models by incorporating a unique structure that includes a cell state and three gates: the forget gate, input gate, and output gate. These components work together to dynamically adjust the network's weights, overcoming the issues of gradient vanishing and exploding. This design allows LSTMs to maintain both long-term and short-term memory effectively [44]. The structure of the LSTM model is illustrated in Fig. 12.

Below, we describe the structure of the three gates in the LSTM model [17].

Forget Gate: The forget gate examines the current time step's input, denoted as x_t, and the output from the previous time step, denoted as h_{t-1}. When $f_t = 0$, the gate discards the read information; conversely, when $f_t = 1$, it preserves the read information. The calculation formula is

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{5}$$

In the equation, σ denotes the sigmoid activation function, W_f represents the weight matrix of the forget gate, and b_f is the bias term.

Input Gate: The input gate determines which new input information to store in the neuron. It starts by creating a candidate cell state $\widetilde{C}_t$, and then updates this state using the input gate i_t. Subsequently, new information is added to the cell state. The specific formula is as follows:

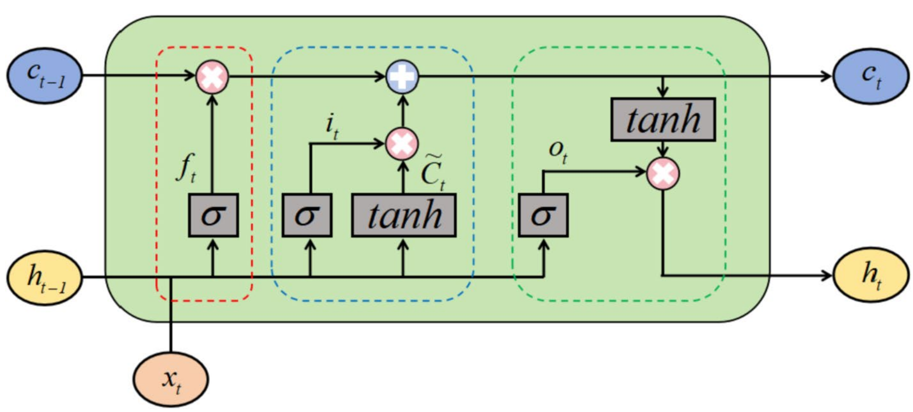

Fig. 12 Architecture of the LSTM network used in this study

$$\widetilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \tag{6}$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \tag{7}$$

$$C_t = f_t \times C_{t-1} + i_t \times \widetilde{C}_t \tag{8}$$

In the equation above, W_c denotes the weight matrix of the cell state, b_c represents the bias term of the cell state, W_i signifies the weight matrix of the input gate, and b_i denotes the bias term of the input gate.

Output Gate: The output gate uses the cell state to determine the final output h_t. It first processes the current input x_t and the previous output h_{t-1}. Then, it multiplies these values by the cell state processed by the tanh layer to produce the ultimate output h_t. The specific formula is as follows:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \tag{9}$$

$$h_t = o_t \times \tanh(C_t) \tag{10}$$

In this formula, W_o represents the weight matrix of the output gate, and b_o denotes the bias term of the output gate.

Details of LSTM network implementation

This study employs an autoencoder LSTM model, which consists of an encoder and a decoder [31]. The encoder converts input vectors of variable length into fixed-length feature vectors that capture the essential attributes of the input. The decoder then reconstructs these fixed-length vectors back into variable-length outputs (as shown in Fig. 13). The final layer consists of a fully connected layer for prediction output. At the end of each batch, the Adam optimization algorithm [45] is employed with mean absolute error (MAE) as the optimization criterion to update network weights and biases. Each batch contains 64 input/output windows, and ReLU activation functions are applied to all LSTM layers [46]. The LSTM autoencoder model was implemented using Python 3 with libraries including PyTorch, NumPy, Pandas, and Scikit-learn.

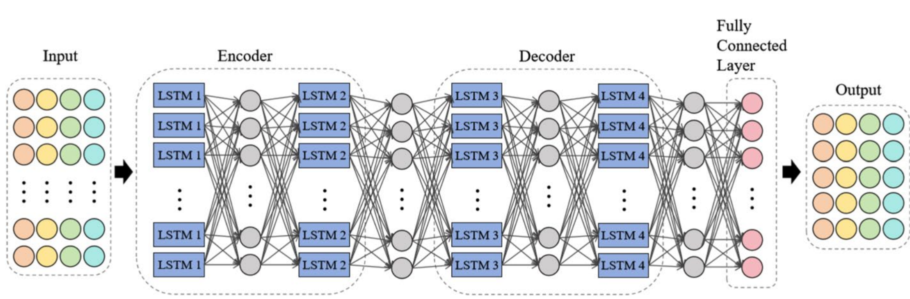

Fig. 13 General structure of the prediction algorithm using the LSTM model with exogenous inputs

Evaluation metrics

To assess network quality, three parameters are considered to quantify the proximity between the predicted variable trajectories $\widehat{y}(Y_1, Y_2, Y_3, Y_4)$ and the actual variable trajectories $y_j(Y_1, Y_2, Y_3, Y_4)$ across the n samples. These calculations are performed after de-standardizing the predicted trajectories (i.e., rescaling them back to their original range). The formula is as follows:

Mean absolute error (MAE):

$$MAE = \frac{1}{n}\sum\nolimits_{j=1}^{n} \left| y_j - \widehat{y}_j \right| \tag{11}$$

Mean squared error (MSE):

$$MSE = \frac{1}{n}\sum\nolimits_{j=1}^{n} \left(y_j - \widehat{y}_j \right)^2 \tag{12}$$

The correlation coefficient (CC) is given as

$$P = \frac{cov(y, \widehat{y})}{std\left(y\right) * std\left(\widehat{y}\right)} \tag{13}$$

where $std()$ is the standard deviation and $cov(y, \widehat{y})$ is the covariance between variables y and $\widehat{y}$.

These metrics are used to evaluate and compare the performance of the network we implemented, and the results were presented in "Results" section.

Abbreviations

LSTM	Long–short-term memory
CNN	Convolutional neural network
MAE	Mean absolute error
MSE	Mean squared error
CC	Correlation coefficient
LElbow	Left elbow
RElbow	Right elbow
LKnee	Left knee
RKnee	Right knee
$\overline{ES}$	The vectors in elbow–wrist direction
$\overline{EW}$	The vectors in elbow–shoulder direction

Acknowledgements
We much appreciate the volunteers for participating in this study.

Author contributions
QX and WH designed the work. YL collected the data. JM analyzed the data. QX and XW interpreted the data. JM and QX drafted the manuscript. YC and NX helped to create the final report.

Funding
This work was supported by the National Natural Science Foundation of China (32460238), and the Natural Science Foundation of Jiangxi Province (20232BAB206134).

Availability of data and materials
The data sets generated and/or analyzed during the current study are not publicly available due to clinical policy but are available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate
The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the ethics committee of the Children's Hospital of Chongqing Medical University. All participants completed informed consent before participation in the protocol.

Competing interests
The authors declare no competing interests.

Received: 20 October 2024 Accepted: 22 February 2025
Published online: 02 April 2025

References

1. Xiong QL, Wu XY, Liu Y, Zhang CX, Hou WS. Measurement and analysis of human infant crawling for rehabilitation: a narrative review. Front Neurol. 2021;12: 731374.
2. McEwan MH, Dihoff RE, Brosvic GM. Early infant crawling experience is reflected in later motor skill development. Percept Mot Skills. 1991;72(1):75–9.
3. Cole WG, Vereijken B, Young JW, Robinson SR, Adolph KE. Use it or lose it? Effects of age, experience, and disuse on crawling. Dev Psychobiol. 2019;61(1):29–42.
4. Crouchman M. The effects of babywalkers on early locomotor development. Dev Med Child Neurol. 1986;28(6):757–61.
5. Yan B, Ying GM. Effect of crawling training on the cognitive function of children with cerebral palsy. Int J Rehabilit Res. 2022;45:184–8.
6. Science DoE, Technology UoS, Technology of China H, Anhui, P.R. China., Science DoE, Technology UoS, Technology of China H, Anhui, P.R. China., Science DoE, Technology UoS, Technology of China H, Anhui, P.R. China., Science DoE *et al*: Human hands-and-knees crawling movement analysis based on time-varying synergy and synchronous synergy theories. Mathe Biosci Eng 2019, 16(4):2492–2513.
7. Chengxiang L, Xiang C, Xu Z, Xun C, De W. Muscle synergy analysis of eight inter-limb coordination modes during human hands-knees crawling movement. Front Neurosci. 2023;17:1135646–1135646.
8. Bell MA, Fox NA. Crawling experience is related to changes in cortical organization during infancy: evidence from EEG coherence. Dev Psychobiol. 1996;29(7):551–61.
9. Yi L, Zhang LF. The effect of crawling training on lower limb function in stroke patients with hemiplegia. China Prac Med. 2013;8(22):261–2.
10. Your total body workout https://www.fitcrawl.com.au/]
11. Ghazi MA, Nash MD, Fagg AH, Ding L, Kolobe THA, Miller DP. Novel assistive device for teaching crawling skills to infants. In: *Field and Service Robotics: Results of the 10th International Conference.* edn. Edited by Wettergreen DS, Barfoot TD. Cham: Springer International Publishing; 2016: 593–605.
12. Jiang JG, Wang CC, Zhang WJ. Design and analysis of a parallel-driven rehabilitation robot for children with cerebral palsy. Mechan Eng. 2020;12:1–3.
13. Lindemann B, Müller T, Vietz H, Jazdi N, Weyrich M. A survey on long short-term memory networks for time series prediction. Procedia CIRP. 2021;99:650–5.
14. Kader NIA, Yusof UK, Khalid MNA, Husain NRN: A Review of long short-term memory approach for time series analysis and forecasting. In: *Proceedings of the 2nd International Conference on Emerging Technologies and Intelligent Systems: 2023// 2023; Cham*: Springer International Publishing; 2023: 12–21.
15. Wen X, Li W. Time series prediction based on LSTM-attention-LSTM model. IEEE Access. 2023;11:48322–31.
16. Tamilselvi C, Paul RK, Yeasin M, Paul AK. Novel wavelet-LSTM approach for time series prediction. Neural Comput Applicat. 2024. https://doi.org/10.1007/s00521-024-10561-z.
17. Goodfellow I, Bengio T, Courville A. Deep learning. 2016.
18. Liu DX, Wu X, Wang C, Chen C. Gait trajectory prediction for lower-limb exoskeleton based on Deep Spatial-Temporal Model (DSTM). In: *2017 2nd International Conference on Advanced Robotics and Mechatronics (ICARM): 27–31 Aug. 2017. 2017*; 2017: 564–569.
19. Zaroug A, Lai DTH, Mudie K, Begg R. Lower limb kinematics trajectory prediction using long short-term memory neural networks. Front Bioeng Biotechnol. 2020;8:362.
20. Su B, Gutierrez-Farewik EM. Gait trajectory and gait phase prediction based on an LSTM network. Sensors. 2020;20(24):7127.
21. Hernandez V, Dadkhah D, Babakeshizadeh V, Kulić D. Lower body kinematics estimation from wearable sensors for walking and running: a deep learning approach. Gait Posture. 2021;83:185–93.
22. Jia L, Ai Q, Meng W, Liu Q, Xie SQ: Individualized gait trajectory prediction based on fusion lstm networks for robotic rehabilitation training. In: *2021 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM): 12–16 July 2021. 2021*; 2021: 988–993.
23. Zaroug A, Garofolini A, Lai DTH, Mudie K, Begg R. Prediction of gait trajectories based on the long short term memory neural networks. PLoS ONE. 2021;16(8): e0255597.
24. Zhu C, Liu Q, Meng W, Ai Q, Xie SQ: An Attention-Based CNN-LSTM Model with Limb Synergy for Joint Angles Prediction. In: *2021 IEEE/ASME International Conference on Advanced Intelligent Mechatronics (AIM).* 12–16 July 20*21 2021*; 2021: 747–752.
25. Challa SK, Kumar A, Semwal VB, Dua N. An Optimized-LSTM and RGB-D sensor-based human gait trajectory generator for bipedal robot walking. IEEE Sens J. 2022;22(24):24352–63.
26. Semwal VB, Jain R, Maheshwari P, Khatwani S. Gait reference trajectory generation at different walking speeds using LSTM and CNN. Multimedia Tools Applicat. 2023;82(21):33401–19.
27. Romero-Sorozábal P, Delgado-Oleas G, Laudanski AF, Gutiérrez A, Rocon E. Novel methods for personalized gait assistance: three-dimensional trajectory prediction based on regression and LSTM models. Biomimetics. 2024;9(6):352.
28. Tanghe K, Groote FD, Lefeber D, Schutter JD, Aertbeliën E. Gait trajectory and event prediction from state estimation for exoskeletons during gait. IEEE Trans Neural Syst Rehabil Eng. 2020;28(1):211–20.

29. Elliott G, Marecki A, Herr H. Design of a clutch-spring knee exoskeleton for running. J Med Devices. 2014;8: 031002.
30. Zhang J, Fiers P, Witte KA, Jackson RW, Poggensee KL, Atkeson CG, Collins SH. Human-in-the-loop optimization of exoskeleton assistance during walking. Science. 2017;356(6344):1280–4.
31. Ding Y, Kim M, Kuindersma S, Walsh CJ. Human-in-the-loop optimization of hip assistance with a soft exosuit during walking. Sci Robot. 2018;3(15):5438.
32. Zaroug A, Proud JK, Lai DTH, Mudie K, Billing D, Begg R. Overview of computational intelligence (CI) techniques for powered exoskeletons. In: *Computational Intelligence in Sensor Networks.* edn. Edited by Mishra BB, Dehuri S, Panigrahi BK, Nayak AK, Mishra BSP, Das H. Berlin, Heidelberg: Springer Berlin Heidelberg; 2019: 353–383.
33. Begg R, Kamruzzaman J. Neural networks for detection and classification of walking pattern changes due to ageing. Australas Phys Eng Sci Med. 2006;29(2):188–95.
34. Begg R, Best R, Dell'Oro L, Taylor S. Minimum foot clearance during walking: strategies for the minimisation of trip-related falls. Gait Posture. 2007;25(2):191–8.
35. Nait Aicha A, Englebienne G, van Schooten KS, Pijnappels M, Kröse B. Deep learning to predict falls in older adults based on daily-life trunk accelerometry. Sensors. 2018;18(5):1654.
36. Hemmatpour M, Ferrero R, Montrucchio B, Rebaudengo M, Piccinno A. A review on fall prediction and prevention system for personal devices: evaluation and experimental results. Adv Hum Comput Interact. 2019;2019:12.
37. Naghavi N, Wade E. Towards real-time prediction of freezing of gait in patients with parkinson's disease: a novel deep one-class classifier. IEEE J Biomed Health Inform. 2022;26(4):1726–36.
38. Banos O, Galvez JM, Damas M, Pomares H, Rojas I. Window size impact in human activity recognition. Sensors. 2014;14(4):6474–99.
39. Kolaghassi R, Al-Hares MK, Marcelli G, Sirlantzis K. Performance of deep learning models in forecasting gait trajectories of children with neurological disorders. Sensors. 2022;22(8):2969.
40. Graves A, Schmidhuber J. Framewise phoneme classification with bidirectional LSTM networks. In: *Proceedings 2005 IEEE International Joint Conference on Neural Networks, 2005: 31 July-4 Aug. 2005 2005*; 2005: 2047–2052 vol. 2044.
41. Graves A. Generating sequences with recurrent neural networks. 2013, abs/1308.0850.
42. Xiong QL, Hou WS, Xiao N, Chen YX, Yao J, Zheng XL, Liu Y, Wu XY. Motor skill development alters kinematics and co-activation between flexors and extensors of limbs in human infant crawling. IEEE Transact Neural Syst And Rehabilitat Engineer. 2018;26(4):780–7.
43. Zhang L, Deng CF, Liu Y, Chen L, Xiao N, Zhai SJ, Hou WS, Chen YX, Wu XY. Impacts of motor developmental delay on the inter-joint coordination using kinematic synergies of joint angles during infant crawling. IEEE Transact On Neural Syst And Rehabilitat Eng. 2022;30:1664–74.
44. Hochreiter S, Schmidhuber J. Long short-term memory. Neural Comput. 1997;9(8):1735–80.
45. Kingma DP, Ba J. Adam: a method for stochastic optimization. 2014, abs/1412.6980.
46. Nair V, Hinton GE. Rectified Linear Units Improve Restricted Boltzmann Machines. In: International conference on machine learning. 2010; 2010.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.